Generalized Boltzmann factors and the maximum entropy principle

Rudolf Hanel[1,2*] and Stefan Thurner[2†]
[1] *Vision Lab; University of Antwerp; Groenenborgerlaan 171; B-2020 Antwerp, Belgium*
[2] *Complex Systems Research Group HNO; Medical University of Vienna; Währinger Gürtel 18-20; A-1090 Vienna; Austria*

We generalize the usual exponential Boltzmann factor to any reasonable and potentially observable distribution function, $B(E)$. By defining generalized logarithms Λ as inverses of these distribution functions, we are led to a generalization of the classical Boltzmann-Gibbs entropy, $S_{BG} = -\int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ \log B(\epsilon)$ to the expression $S \equiv -\int d\epsilon\ \omega(\epsilon) \int_0^{B(\epsilon)} dx\ \Lambda(x)$, which contains the classical entropy as a special case. We demonstrate that this entropy has two important features: First, it describes the correct thermodynamic relations of the system, and second, the observed distributions are straight forward solutions to the Jaynes maximum entropy principle with the ordinary (not escort!) constraints. Tsallis entropy is recovered as a further special case.

PACS numbers: 05.20.-y, 89.75.-k, 05.70.-a, 05.90.+m

INTRODUCTION

It has been realized, that many statistical systems in nature can not be satisfactorially described by naive or straight forward application of Boltzmann-Gibbs statistical mechanics. In contrast to ergodic, separable, locally and weakly interacting systems, these systems are *complex* systems whose characteristic distributions often are of power-law type. Due to the existence of strong correlations between its elements complex systems often violate ergodicity and are prepared in states at the *edge of chaos*, i.e. they exhibit weak sensitivity to initial conditions. Further, complex systems are mostly not separable in the sense, that probabilities for finding a system in a given state factorize into single particle probabilities and as a consequence, renders these systems not treatable with Boltzmann single particle entropies [1]. However, it is evident that Gibbs entropies can in principle take into account any correlations in a given system, as the full Hamiltonian H, including potential terms, enters. Since in the following we will be only concerned about measurable quantities in statistical systems we will take the Gibbs entropy as a starting ground

$$S_G = -\int d\Gamma\ B\left(H(\Gamma)\right)\ \log\left(B\left(H(\Gamma)\right)\right) \quad , \qquad (1)$$

where Γ are the phase space variables, and B is the Boltzmann factor, which usually reads, $B(H) \sim \exp(-\beta H)$, for the canonical distribution. It is interesting to note that the exponential form of the Boltzmann factor is not a priori dictated by classical statistical mechanics, but that much of classical statistical mechanics is built upon this special form of the Boltzmann factor, as argued e.g. in [2].

Classical statistical mechanics was designed for systems with short- (or zero-) range interactions, such as gas-dynamics. The exponential was found to be the natural choice in countless systems. However, for extending the concept of statistical mechanics to complex systems, which are characterized by fundamentally different distribution functions, it seems natural to allow generalizations of the Boltzmann factor. What is the Boltzmann factor? What are the minimum requirements and restrictions to call some function a Boltzmann factor?

The normalized Boltzmann factor is a probability to encounter a particular state in the bath system, representing the hidden physical influences the observable ensemble of properties are subject to and thus closely relates to experiment. In the canonical ensemble the density of states with energy E_1 are given by

$$\rho(E_1) = \omega_1(E_1)\omega_2(E - E_1)Z^{-1} \quad , \qquad (2)$$

where ω_1 is the subjective microcanonical density, i.e. the multiplicity of states in the ensemble of observable properties, and ω_2 is the bath density. E is the energy of the total system, which is usually unknown, and Z is the partition function. Usually, the normalized $\omega_2(E - E_1)Z^{-1}$ is identified with the Boltzmann factor. However, in this form it explicitly depends on the total system energy E. This total energy should be factored out into a multiplicative factor since measured quantities should not depend on E. This factor will be canceled by Z, which is of course E dependent. If the Boltzmann factor is taken as an exponential, this separation is trivial. Another approach is to ask which classes of Boltzmann factors allow for such a factorization. The answer was given in [3], showing by a mathematical argument, that the most general Boltzmann factors which allow for an E separation are of so-called q-exponential type.

In the following, we start by exploring a most general form of the Boltzmann factor, compatible with the requirements of normalizability, monotonicity and the possibility of E separation. We do not fix the specific form of this factor which (in principle) can be determined from measurements. We ask whether one can construct a theoretical framework where data, i.e. the

measured distribution serves as a starting point to construct an entropy which is consistent with both, the correct thermodynamic relations and the Jaynes maximum entropy principle [4]. According to this modification of logics it is sensible in a first step to modify or deform the log in Eq. (1) to a generalized logarithm Λ. The concept of deforming logarithms and thus modifying the form of entropy in order to accommodate a large body of experimental data from complex systems is not new [2, 5, 6, 7, 8, 9, 10, 11]. An axiomatic definition of generalized logarithmic and exponential functions Λ and $\mathcal{E}$ has been given in [12] where also the concept of dual logarithms of the form $\Lambda^*(x) \equiv -\Lambda(1/x)$ has first been introduced. An algebraization of the deformed concept, i.e. $x \otimes y = \mathcal{E}(\Lambda(x)+\Lambda(y))$, and $x \oplus y = \Lambda(\mathcal{E}(x)\mathcal{E}(y))$, has been given in [2], where this structure has been exploited in the context of special relativistic mechanics. In [13] a constrained variational principle has then been utilized with respect to trace-form entropies deriving a family of three-parameter deformed logarithms $\log_{(\kappa,r,\zeta)}$, being the most general of its kind so far, containing – to our best knowledge – all possible logarithms that are compatible with the standard variational principle $\delta\tilde{G} = 0$, with the usual functional

$$\begin{aligned} \tilde{G} = \tilde{S}_G[B] \;\; &- \;\; \beta \int d\epsilon \; \omega\left(\epsilon\right) B\left(\epsilon\right) \left(\epsilon - U\right) \\ &- \;\; \gamma \left(\int d\epsilon \; \omega\left(\epsilon\right) B(\epsilon) - 1 \right) \quad , \end{aligned} \tag{3}$$

with the generalized Gibbs entropy

$$\tilde{S}_G[B] = -\int d\epsilon \; \omega\left(\epsilon\right) B\left(\epsilon\right) \; \Lambda\left(B\left(\epsilon\right)\right) \quad , \tag{4}$$

where U is the measured average energy, $\omega(\epsilon)$ is the multiplicity, β is the usual inverse temperature, and γ is the Lagrange parameter for normalizability.

The novel logics of this paper is that we start from a measured distribution, the Boltzmann factor, which is not necessarily of standard exponential form. We want to keep the intuition of the origin of the Boltzmann factor as the adequately normalized contributions of the bath, i.e. we require $\rho(E) = \omega(E)B(E)$, where ω is the multiplicity of the energy state in the observable system and represents our knowledge about the experimental device we observe in order to retrieve data. In principle, ω can be known which makes the Boltzmann factor $B(E) = \rho(E)/\omega(E)$ factor indirectly measurable. To keep close formal contact with usual statistical physics, we represent the measured Boltzmann factor by replacing the usual exponential function by some function $\mathcal{E}$, i.e.

$$\exp(-\beta(E-U) - \tilde{\gamma}) \rightarrow \mathcal{E}(-\beta(E-U) - \tilde{\gamma}) \quad , \tag{5}$$

where $\tilde{\gamma}$ is the normalization constant. We then construct an entropy such that two requirements are strictly fulfilled: First, the entropy leads to the correct thermodynamics of the system, and second, the Jaynes variational principle holds.

THE GENERALIZED BOLTZMANN FACTOR

Let us start by listing three "axioms" containing the intuitively clear minimum requirements for a Boltzmann factor B,

1. B is monotonous and positive.
2. B can be normalized, i.e. $\int d\epsilon \, \omega_1(\epsilon) \, B(\epsilon) = 1$.
3. B must not explicitly depend on the total system energy. It must be possible that the E term in the argument of $\omega_2(E - E_1)$ can be factored out, i.e., $\omega_2(E - E_1) = F(E - E^*) \; B(E_1 - E^*)$, where the normalized version of B we shall call a Boltzmann factor. F is some function, and E^* some reference energy, e.g. the equilibrium energy. In [3] and [14] an explicit program was shown how this separation is uniquely obtained.

We thus write a Boltzmann factor which fulfills all requirements

$$B(H) \equiv \mathcal{E}(-\beta(H-U) - \tilde{\gamma}) \quad , \tag{6}$$

where $\tilde{\gamma}$ is the normalization constant (partition function), U and β being the measured average energy and inverse temperature, respectively. Monotonicity and positivity are assumed to be properties of the generalized exponential functions $\mathcal{E}$, which then implies the existence of inverse functions, the associated generalized logarithms $\Lambda = \mathcal{E}^{-1}$. From a generalized logarithm Λ and its dual $(\Lambda^*(x) \equiv -\Lambda(x^{-1}))$ one assumes the usual properties,

$$\begin{aligned} &\Lambda : \mathbf{R}^+ \rightarrow \mathbf{R} \\ &\Lambda(1) = 0 \quad , \quad \Lambda'(1) = 1 \quad , \quad \Lambda' > 0 \\ &\Lambda'' < 0 \qquad \text{(convexity)} \quad , \end{aligned} \tag{7}$$

implying analogous properties for the generalized exponential function.

Now, with any representative of the above allowed generalized Boltzmann factor B and its associated logarithm Λ let us in a first step generalize Gibbs entropy Eq. (1), (same as Eq. (4)),

$$\tilde{S}_G \equiv -\int d\Gamma B\left(H(\Gamma)\right) \Lambda\left(B\left(H(\Gamma)\right)\right) \quad , \tag{8}$$

and compute the Gibbs entropy as follows

$$\begin{aligned} \tilde{S}_G \;&=\; -\int d\Gamma \;\, B\left(H\right) \Lambda\left(B\left(H\right)\right) \\ &=\; -\int d\epsilon \int d\Gamma \delta\left(\epsilon - H\right) B\left(\epsilon\right) \Lambda\left(B\left(\epsilon\right)\right) \\ &=\; \int d\epsilon \; \omega_H(\epsilon) \; \mathcal{E}\left(-\beta(\epsilon - U) - \tilde{\gamma}\right) \left(\beta(\epsilon - U) + \tilde{\gamma}\right) , \end{aligned} \tag{9}$$

where $\omega_H(E) \equiv \int d\Gamma \delta(E - H)$ is the microcanonic multiplicity factor for the energy E which represents the observable system. As a shorthand notation we will write Eq. (9) as in Eq. (4), $\tilde{S}_G = \int d\epsilon \; \omega(\epsilon) \; B(\epsilon) \; \Lambda(B(\epsilon))$, with

$B(E) = \mathcal{E}(-\beta(E-U) - \tilde{\gamma})$. With the definition of the expectation value

$$\langle f \rangle \equiv \int d\epsilon\, f(\epsilon)\, \omega_H(\epsilon)\, \mathcal{E}\left(-\beta(\epsilon - U) - \tilde{\gamma}\right) \quad , \qquad (10)$$

it becomes obvious that the normalization constant $\tilde{\gamma}$ has to be chosen such that

$$\int d\epsilon\ \omega_H(\epsilon)\ \mathcal{E}\left(-\beta(\epsilon - U) - \tilde{\gamma}\right) = 1 \quad . \qquad (11)$$

Using this and specifying $\langle \epsilon \rangle = U$, we get $\tilde{S}_G = \tilde{\gamma}$. We drop the subscript H in the following. Looking at $\tilde{S}_G$ for $\beta = 0$, implies that $B(E) = Z^{-1} = \text{const}$, for $Z = \int d\epsilon\ \omega(\epsilon)$, and therefore $\tilde{S}_G = -\int d\epsilon\ \omega(\epsilon) Z^{-1} \Lambda(Z^{-1}) = -\Lambda(Z^{-1})$. Thus one identifies

$$\tilde{S}_G = \tilde{\gamma} = -\Lambda(Z^{-1}) = \Lambda^*(Z) \quad . \qquad (12)$$

Note, that to get a finite Z it is necessary to understand the integral $\int d\epsilon \omega(\epsilon)$, in the limits $E_1 = 0$, and $E_2 = E_{\max}$, where $E_{\max}$ is the largest energy of the observable system. Such regularizations are of course implicitly present under all experimental circumstances. If we wish this relation to hold for all β it is interesting to observe that the partition function Z also has to be defined in a deformed way, i.e. using the definition of the deformed product $x \otimes y = \mathcal{E}(\Lambda(x) + \Lambda(y))$, analogous to [2]. The renormalization condition can then be recast into the form

$$B(H) = \left(\frac{1}{Z}\right) \otimes \mathcal{E}\left(-\beta\left(H(\Gamma) - U\right)\right) \quad , \qquad (13)$$

which becomes the defining equation for the generalized partition function Z. In this sense the generalization of Boltzmann factors naturally involves dual logarithms, whose occurence has been noted recently in the context of generalized entropies [6, 8, 9, 10]. This is of course just of relevance for non self-dual logs, examples of which include the q-logarithm ($\log^*_q(p) = \log_{2-q}(p)$) and the Abe-log [7].

THE VARIATIONAL PRINCIPLE

Using the standard variational principle Eq. (3) on the basis of the generalized entropy given in Eq. (9) (with the usual constraints!), the only possible choice for Λ is the ordinary log. To see this, variation of Eq. (3) yields

$$\frac{d}{dB} B\Lambda(B) = -\gamma - \beta(E-U) \quad . \qquad (14)$$

By substituting $B = \mathcal{E}\left(-\beta(\epsilon - U) - \tilde{\gamma}\right)$, it is clear that the only solution to this is $\Lambda(B) = \log(B)$, and $\mathcal{E}$ can thus only be the ordinary exponential Boltzmann factor. This is unsatisfactory.

The problem arises because for any generalized Λ other than the ordinary log there exists a non-trivial extra term, $B\Lambda'(B)$, in Eq. (14). In order to cancel this term we suggest to further generalize the generalized logarithm $\Lambda(B)$ to a *functional* in the following way,

$$\Lambda(B) \to \bar{\Lambda}[B] \equiv \Lambda(B) - \eta[B] \quad , \qquad (15)$$

where we use $[B]$ to indicate functional dependence on B. By substituting Λ by $\bar{\Lambda}$ in Eq. (4), we obtain the entropy

$$S[B] \equiv \tilde{S}_G[B] + \eta[B] \quad , \qquad (16)$$

where we have used that η is a constant with respect to ϵ-integration and the normalization condition (11). Now the idea is that after variation with respect to B, the additional term $\frac{\delta}{\delta B}\eta[B]$, can be used to cancel the term $-\omega(E)B(E)\frac{d}{dB}\Lambda(B(E))$. The corresponding condition, $\frac{\delta}{\delta B}\eta[B] = \omega(E)B(E)\frac{d}{dB}\Lambda(B(E))$, implies the form of η

$$\eta[B] = \int d\epsilon\ \omega(\epsilon) \int_0^{B(\epsilon)} dx\ \Lambda'(x)x + c \quad , \qquad (17)$$

Let us substitute this into Eq. (16) to get

$$\begin{aligned} S[B] &= \eta[B] - \int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ \Lambda(B(\epsilon)) \\ &= -\int d\epsilon\ \omega(\epsilon) \int_0^{B(\epsilon)} dx\ \Lambda(x) + \bar{c} \quad , \end{aligned} \qquad (18)$$

with $\bar{c}$ an integration constant which is only different from c, iff $\lim_{x\to 0} x\Lambda(x) \neq 0$. Note immediately that the classical entropy is a special case of Eq. (18), i.e. taking $\Lambda(x) = \log(x)$, yields the Boltzmann entropy modulo an irrelevant additive constant, $S[B] = -\int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ \log(B(\epsilon)) + \bar{c} + 1$.

It can now easily be checked that this entropy Eq.(18), in combination with the standard maximum entropy principle under the *usual* constraints, yields the measured distributions B. Let us define

$$\begin{aligned} G = S[B] &- \beta \int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ (\epsilon - U) \\ &- \gamma\left(\int d\epsilon\ \omega(\epsilon)\ B(\epsilon) - 1\right) \quad , \end{aligned} \qquad (19)$$

and vary with respect to B, to get

$$\begin{aligned} \frac{\delta}{\delta B} G &= \omega(\epsilon)B(E)\Lambda'(B(E)) - \frac{d}{dB}\omega(E)B(E)\Lambda(B(E)) \\ &- \omega(E)\gamma - \omega(E)\beta(E-U) = 0 \quad , \end{aligned} \qquad (20)$$

or equivalently, $\Lambda(B(E)) = -\gamma - \beta(E-U)$. Using that $\mathcal{E}$ is the functional inverse of Λ, the correct generalized Boltzmann factor, $B(E) = \mathcal{E}(-\beta(E-U) - \gamma)$, is recovered.

THERMODYNAMICS

To show that the proposed entropy of Eq. (18) is fully consistent with the expected thermodynamic relations, differentiate Eq. (18) with respect to U and get

$$\frac{\partial}{\partial U} S[B] = \beta \quad . \qquad (21)$$

Note, that the thermodynamics here is simply $dU = TdS$, since no further assumptions have been made on other measurements neither in terms of thermodynamic potentials (e.g. $-PdV$ or $-\mu dN$) nor other (experimentally controllable) macro-state variables.

Finally, if one wants to write the proposed entropy Eq. (18) in a form that is suggested by the classical Gibbs form one can, by defining L, of course write

$$\begin{aligned} S[B] &= -\int d\epsilon\ \omega(\epsilon) \int_0^{B(\epsilon)} dx\ \Lambda(x) \\ &\equiv -\int d\epsilon\ \omega(\epsilon)\ B(\epsilon) L(B(\epsilon)) \quad , \end{aligned} \tag{22}$$

which implies the relation

$$L(a) = \frac{1}{a} \int_0^a dx\ \Lambda(x) \quad . \tag{23}$$

It is maybe interesting to note that L is nothing but the mean value of the Λ. Of course, in general L is not an inverse of B.

EXAMPLES

Example: Classical Boltzmann distributions. If the experimentally measured tail of a distribution is of Boltzmann type, $B(E) \sim \exp(-\beta E)$, then $\Lambda(B) \sim \log(B)$, and by using Eq. (23), $L(B) = \frac{1}{B}(B\log(B) - B)$, which when put into Eq. (22), yields the Boltzmann entropy, $S[B] = -\int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ \log(B(\epsilon)) + 1$.

Example: Asymptotic power-law distributions. If an experimental distribution of a q-exponential is observed as frequently done in complex systems, i.e. $B(E) = [1-(1-q)E]^{\frac{1}{1-q}}$. Thus the generalized logarithm is the so-called q-log, $\Lambda(B) = \log_q(B) \equiv \frac{B^{1-q}-1}{1-q}$. Inserting as before gives the Tsallis entropy [5, 6] times a factor,

$$S[B] = -\frac{1}{2-q} \int d\epsilon\ \omega(\epsilon)\ B(\epsilon)\ \log_q(B(\epsilon)) + \frac{1}{2-q} \quad , \tag{24}$$

where we require $q < 2$. The factor can in principle be absorbed into a transformation of β and γ. At this point it is also obvious that in the case of power law distributions the question of normalizability can become an issue. Notice, however, that since not B but $\rho = \omega B$ has to be normalizable an implicit regularization is provided by the maximal energy $E_{\max}$ that the observable system, represented by ω, can assume.

CONCLUSION

We started by relaxing the restriction that the Boltzmann factor has to be of exponential form, to allow other types of observed distributions, $B(E)$, as well. By doing so we introduce corresponding generalized logarithms, Λ (as inverses of B), and suggest to construct the entropy of systems leading to non-exponential distributions, as $S = -\int d\epsilon\ \omega(\epsilon) \int_0^{B(\epsilon)} dx\ \Lambda(x)$. This is nothing but replacing the $p\log p$ term in the usual entropy by the integral, $\int \Lambda(p)$. Obviously classical Boltzmann-Gibbs entropy is obtained for the special case of $\Lambda(x) = \log(x)$. We demonstrate that this entropy leads to the correct thermodynamics of the system, and the observed distribution functions are derived naturally from the maximum entropy principle with the usual constraints. Further we show that this entropy can be written as a standard generalized Gibbs entropy ($\int B\Lambda B$) with adding a constant which is functionally dependent on the measured distribution [15]. This term somehow captures numbers of states in the system, which may depend on parameters like temperature. The functional form of measured distributions, which is a kind of knowledge about the system, is thus naturally fed into the definition of the entropy of the system.

A further detail in the proposed entropy definition is that it does not contain any additional parameters, once the distribution is known. Once given the data, there is no more freedom of choice of the generalized logarithms, nor of the functional form of the constant.

We thank J. Lutsko for a number of important suggestions.

* Electronic address: Rudolf.Hanel@ua.ac.be
† Electronic address: thurner@univie.ac.at

[1] L. Boltzmann, Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften **66**, 275-370 (1872).
[2] G. Kaniadakis, Phys. Rev. E **66**, 056125 (2002).
[3] R. Hanel and S. Thurner, Physica A **351**, 260 (2005).
[4] E.T. Jaynes, Phys. Rev. **106**, 620-630 (1957).
[5] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[6] C. Tsallis, M. Gell-Mann and Y. Sato, Proc. Natl. Acad. Sc. USA **102**, 15377 (2005).
[7] S. Abe, Phys. Lett. A **224**, 326 (1997).
[8] T. Wada and A.M. Scarfone, Phys. Lett. A **335**, 351 (2005).
[9] J. Naudts, Chaos, Solitons and Fractals **13**, 445 (2002).
[10] E.P. Borges, cond-mat/0304545.
[11] T.M. Rocha Filho, A. Figueiredo and M.A. Amato, Phys. Rev. Lett. **95**, 190601 (2005).
[12] J. Naudts, Physica A **316** (2002) 323.
[13] G. Kaniadakis, Phys. Rev. E **72**, 036108 (2005).
[14] R. Hanel and S. Thurner, Physica A **365**, 162-166 (2006).
[15] The generalization from a function to a functional was suggested by J. Lutsko (private communication).